DOCUMENT RESUME

ED 398 240 TM 024 740

AUTHOR Scriven, Michael
TITLE Student Ratings Offer Useful Input to Teacher Evaluations. ERIC/AE Digest.
INSTITUTION ERIC Clearinghouse on Assessment and Evaluation, Washington, DC.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO EDO-TM-95-8
PUB DATE Oct 95
CONTRACT RR93002002
NOTE 4p.; Condensed from "Using Student Ratings in Teacher Evaluation."
AVAILABLE FROM ERIC Clearinghouse on Assessment and Evaluation, The Catholic University of America, Department of Education, O'Boyle Hall, Washington, DC 20064 (free).
PUB TYPE Information Analyses - ERIC Clearinghouse Products (071)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Faculty; *College Students; Data Collection; *Decision Making; Higher Education; *Instructional Improvement; *Student Evaluation of Teacher Performance; *Teacher Effectiveness; *Validity
IDENTIFIERS ERIC Digests

ABSTRACT

Student ratings of instruction are widely used as a basis for personnel decisions and faculty development recommendations. This digest discusses concerns about the validity of student ratings and presents a case for their use in teacher evaluation. There are several strong arguments for using student ratings to evaluate teachers. Students are in a unique position to rate their own increased knowledge and comprehension, and they are able to rate aspects of instructor competence and attitude. Nine possible lines of argument for the validity of student ratings are summarized, but it is noted that student ratings become invalid if the rating form used is not appropriate for the data collection required. Several points in form design are described, and errors to be avoided are noted. Student ratings must be considered carefully in the context in which they are given, and they cannot be the only source for teacher evaluation. (Contains 4 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 398 240

STUDENT RATINGS OFFER USEFUL INPUT TO TEACHER EVALUATIONS

ERIC/AE Digest Series EDO-TM-95-8
The Catholic University of America

Michael Scriven

TM024740

BEST COPY AVAILABLE

CUA

ERIC/AE DIGEST

DEPARTMENT OF EDUCATION THE CATHOLIC UNIVERSITY OF AMERICA

EDO-TM-95-8 October 1995

Student Ratings Offer Useful Input to Teacher Evaluations

condensed from "Using Student Ratings in Teacher Evaluation," by Dr. Michael Scriven, Project Director, Teacher Evaluation Models Project, Center for Research on Educational Accountability and Teacher Evaluation (CREATE)

Student ratings add a valuable component to the range of input for the evaluation of teachers. Although many question the validity of such ratings, under certain conditions, results can and should be useful.

Student ratings of instruction are widely used as a basis for personnel decisions and faculty development recommendations in post-secondary education today. This article addresses concerns about their validity and presents a case for the use of student ratings in teacher evaluation. In this discussion, student ratings refer to those in which students are asked to complete a form or write a short free-form evaluation anonymously, either during or immediately after a class period, the final exam, or a session after grades are issued

Concerns

Oftentimes, student rating forms ask many questions about matters that students do not appear to be in any position to judge reliably. In addition, the fact that the overall rating of teaching merit by students is only statistically related to learning gains is a concern if one believes that statistical indicators should not be used to make personnel decisions. Another concern is that the validation studies that are used to justify student ratings use questionable indicators instead of the true criterion. For example, some of them correlate the student ratings with peer ratings of teacher merit instead of with superior learning gains.

Arguments for using student ratings

There are several strong arguments for using student ratings to evaluate teachers. (See box titled "Nine Potential Sources of Validity for Student Ratings of Instruction.") Students are in a unique position to rate their own increased knowledge and comprehension as well as changed motivation toward the subject taught. As students, they are also in a good position to judge such matters as whether tests covered all the material of the course.

Nine Potential Sources of Validity for Student Ratings of Instruction

1. The positive and statistically significant correlation of student ratings with learning gains.
2. The unique position and qualifications of the students in rating their own increased knowledge and comprehension.
3. The unique position of the students in rating changed motivation (a) toward the subject taught; perhaps also (b) toward a career associated with that subject; and perhaps also (c) with respect to a changed general attitude toward further learning in the subject area, or more generally.
4. The unique position of the students in rating observable matters of fact relevant to competent teaching, such as the punctuality of the instructor and the legibility of writing on the board.
5. The unique position of the students in identifying the regular presence of teaching style indicators. Is the teacher enthusiastic; does he or she ask many questions, encourage questions from students, etc.?
6. Relatedly, students are in a good position to judge--although it is not quite a matter of simple observation--such matters as whether tests covered all the material of the course.
7. Students as consumers are likely to be able to report quite reliably to their peers on such matters of interest to them as the cost of the texts, the extent to which attendance is taken and weighted, and whether a great deal of homework is required--considerations that have little or no known bearing on the quality of instruction.
8. Student ratings represent participation in a process often represented as "democratic decisionmaking."
9. The "best available alternative" line of argument.

In addition, students can observe and rate facts (i.e. an instructor's punctuality, the legibility of writing on the board) that are relevant to competent teaching. They can also identify and rate whether the teacher is enthusiastic? does he or she ask many

ERIC Clearinghouse on Assessment and Evaluation

The Catholic University of America, Department of Education, O'Boyle Hall, Washington, DC 20064 202 319-5120

questions? encourage questions from students, etc.?

Evaluation Forms

However, the possible lines of argument (see box above) for the validity of student ratings become invalid if the rating form used is not appropriate for the specific data collection required. Since rating forms vary widely, generalizations about student ratings as a good indicator of learning gains or teacher merit are misleading since they assume there is a common property to all such ratings. Most forms, when used in the most common ways, are invalid as a basis for personnel action. For example, many forms used to make personnel decisions ask questions that may influence the respondent by mentioning extraneous and potentially prejudicial material (i.e., questions about the teacher's personality or the appeal of the subject matter).

Another problem with the use of rating forms for summative evaluation is that many of them ask the wrong global or overall questions. This is important since it is typically these questions on which most personnel decisions are based. Common examples of this kind of mistake include forms that ask for

- comparisons with other teachers,
- whether the respondent would recommend the course to a friend with similar interests, or
- whether "it's one of the best courses" one has had.

Several pragmatic considerations (logistical, political, economic, psychological), which impact form design, are required for validity. These include:

- Form length--if forms are too long students may not fill them in or may skip responses.
- Type of question--forms should include the questions students want answered about the courses they are considering taking, thus avoiding resentment and a lack of willingness to complete the forms; **forms should not include** questions that students suspect will be used to discriminate against them or that are biased towards favorable (or unfavorable) comments.

The validity of student rating forms is also dependent on the context of how and when they are administered. For student rating results to be valid, they must be obtained from properly administered tests, stringently controlled data collection, and thorough analysis of test results. Frequent errors include

- The use of instructors to collect forms rating their own instructional merit.
- Lack of controls over pleas for sympathy or indulgence by the teacher before forms are distributed.
- Inadequate time to complete forms.
- Failing to ensure an acceptable return rate.

To ensure the validity of results, errors in data processing, report design, and interpretation must also be avoided. Common errors include

- The use of averages alone, without regard to the distribution.
- Failure to set up appropriate comparison groups so that the usual tendency for ratings to be higher in graduate professional schools can be taken into account.
- Treating small differences as significant, just because they are statistically significant.
- Using factor analysis without logical/theoretical validation.
- Ignoring ceiling/floor effects.
- Using the ratings as the sole basis for either formative or summative evaluation.

Although student ratings are an important source of data for the evaluation of teaching merit, they should not be the only source. Similarly, student ratings form an essential part of the data for the evaluation of courses, workshops, degree programs, etc., but they cannot carry the entire burden. It is essential to look at the data relating to other dimensions of merit such as needs, demand, opportunities for symbiosis, content, and costs, and estimate their relative importance.

Student ratings must be considered very carefully in the context in which they are given. The educational administrator interested in the improvement of instruction--whether by improving courses themselves, or the performance or the composition of the faculty--and instructors and students with the same interest will benefit from the use of a sound system of student ratings.

Further Reading

Abrami, P.C.(1989). How Should We Use Student Ratings to Evaluate Teaching? *Research in Higher Education*, 30 (2), 221-227.

Abrami, P.C., d'Apollonia, S., &P.A. Cohen (1990). Validity of Student Ratings of Instruction: What We Know and What We Do Not Know. *Journal of Educational Psychology*, 82 (2), 219-231.

L'Hommedieu, R.Menges, R.J. & K.T. Brinko (1990) Methodological Explanations for the Modest Effects of Feedback from Student Ratings. *Journal of Educational Psychology*, 82 (2), 232-241

Scriven, M. (1994) Using Student Ratings in Teacher Evaluation, *Evaluation Perspectives* (Newsletter of The Center for Research on Educational Accountability and Teacher Evaluation, 4(1), 1-4.

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract RR93002002. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**